

SHORT THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (PHD)

**Modification of reinforcement learning processes
in mood disorders in asthma patients**

by Dr. Magdolna Emma Szilasi

Supervisor: Dr. Judit Zsuga



UNIVERSITY OF DEBRECEN
DOCTORAL SCHOOL OF HEALTH SCIENCES

DEBRECEN, 2018

Modification of reinforcement learning processes in mood disorders in asthma patients

By Magdolna Szilasi MD

Supervisor: Judit Zsuga MD, PhD

Doctoral School of Health Sciences (Preventive Medicine and Public Health program)
University of Debrecen

Head of the **Defense Committee:** Róza Ádány, MD, PhD, DSc

Reviewers: Beatrix Bálint, MD, PhD
László Brugós, MD, PhD

Members of the Defense Committee: Zoltán Bártfai, MD, PhD
Miklós Vecsernyés, PharmD, PhD

The PhD Defense takes place at the Lecture Hall of Bldg. A, Department of Internal Medicine, Faculty of Medicine, University of Debrecen at 1 p.m., June 15, 2018 .

1. Introduction.....	5
1.1. General overview.....	5
1.2. Reinforcement learning: Bellman equation.....	7
1.3. The proactive model of the reinforcement learning.....	8
1.4. Association between irisin/BDNF axis and brain reward system.....	9
1.5. The relationship between the reinforcement learning and mood disturbance.....	10
1.6. Mood disturbance, well-known co-morbid disorder of asthma.....	11
1.7. Aims and goals.....	11
2. Materials and methods.....	12
2.1. Conceptualization.....	12
2.2. Study Design and protocol.....	13
2.3. Pulmonary function testing.....	14
2.4. Blood samples.....	14
2.5. Determination of serum irisin and BDNF levels.....	15
2.6. St. George's Respiratory Questionnaire (SGRQ).....	15
2.7. Statistical analysis.....	16
3. Results	18
3.1. The proactive brain builds a model of the environment.....	18
3.2. The orbitofrontal cortex compounds the reward function attribute of the Bellmann equation	19
3.3. Association between serum levels of irisin and BDNF and mood disturbances in asthma patients	21
3.3.1. Patients.....	21
3.3.2. Comparison of patients regarding distress disorder.....	21
3.3.3. Association between SGRQ's Impacts score, serum irisin and BDNF levels.....	21
4. Discussion.....	23
4.1. 'Proactive' use of cue-context congruence.....	23
4.1.1. Specificity of the OFC regarding the reward function	23
4.1.2. Computation of the model-free reward prediction error using input from the OFC.....	24
4.1.3. Updating the world model	25
4.1.4. Clinical implications of our hypothesis	26
4.1.5. The relationship between irisin/BDNF axis and mood in asthma patients.....	26

5. New findings.....	29
6. Summary	30
7. Acknowledgements.....	31

Table of contents

1. Introduction

1.1. General overview

The most significant public health challenge of the 21st century is to combat non-communicable chronic diseases. The main goal of health policy is to avoid premature death and reduce healthcare spendings. The four major behavioral risk factors smoking, alcohol consumption, dietary problems and lack of physical activity are associated with four major disease clusters: cardiovascular diseases, cancer, diabetes, and chronic respiratory diseases. This is described by a number of international policy directives, including a global summary of the Global Action Plan for the Prevention and Control of Non-Communicable Diseases 2013-2020 issued by the World Health Organization (WHO).

Chronic respiratory inflammatory diseases include various types of asthma, which, in addition to the known serious social and economic burdens, are important to mention regarding their outstanding impact on the quality of life of patients. Asthma is a chronic inflammatory disorder of the respiratory tract, resulting in recurrent episodes of wheezing, varying dyspnoea severity, chest tightness and coughing. These symptoms are usually associated with diffuse but varying degrees of airway obstruction. Asthma can not be cured, but proper medication and preventive activities can control the symptoms well. Therefore, the current international directive (GINA 2017) recommends an asthma-control based therapy for continuous follow-up of patients' condition and, consequently, a repeated review of therapy.

A commonly-researched and decisive factor in people with chronic illness is the degree of physical and mental well-being, e.g. quality of life. The emergence of quality of life considerations in medical science was based on the Health Definition of the World Health Organization of 1946, which states that "Health is a state of complete physical, psychological and social well-being and not just a lack of disease or disability".

The disorder greatly influences the quality of life and has an impact on physical condition and everyday physical activity. Among asthma patients various mental illnesses are very common, which in turn contribute to worse asthma control and quality of life. In the event of these mental comorbidities (e.g. depression, generalized anxiety), simultaneous dyspnoea may be a threatening experience for the patient, which greatly impairs the physical symptoms associated with psychic illness, thus worsening them and causing panic attacks.

Recently, narrowed and rigid contextual learning was proposed as a potential common neurobehavioral mechanism that may underlie distress disorder. Contextual learning is a basic mechanism involved in reinforcement (reward) learning and motivation to organize cues and their respective contexts (including rewards) into context frames based on their statistical regularities. These will serve as the starting point for making forward looking simulations to maximize the sum of future rewards and govern motivated behavior. Characteristic symptoms such as depressive rumination and anxiety have been directly linked to the alteration of contextual learning in distress disorder.

One possible molecular pathway that may link affective constructs (anxiety, rumination) to reinforcement learning is the irisin-brain-derived neurotrophic factor (BDNF) axis, on basis of its potential role in reinforcement learning, given its modulatory effect in structures related to the generation of context frames inherent of reinforcement learning, on the other hand, it may be involved in altered mood, as raised by the neurotrophic theory of depression.

The change in behavior is fundamentally determined by motivation and reward two constructs closely linked to the paradigm of reinforcement learning and the reward-related learning processes. Central to these processes is the mesocortico-limbic system. In the background of the previously mentioned risk behaviors, the dysfunction of the mesocortico-limbic system has been proposed previously, as a pathological factor, raising its central role in the prevention and progression of non-communicable diseases. Based on these clinical observations, it is essential to identify the neural structures underlying behavioral change, to investigate their functions, and to identify

neurobiological processes that will allow the identification of new diagnostic markers and therapeutic targets.

1.2. Reinforcement learning: Bellman equation

Reinforcement learning is a fundamental form of learning where learning is governed by the rewarding value of a stimulus or action. The Bellman equation is a central theorem in reinforcement learning, it defines the value of a given state as the sum of the immediate reward received upon entering a state and the discounted value of future states that may be obtained starting from the current state.

In reinforcement learning the agent is interacting with the environment, which is reflected in the individual's actions. The agent is constantly receiving feedback from his environment, choosing an action accordingly, then getting a reward for the action he chooses and getting into a new state. According to the Bellman equation, the value of a state is the discounted sum of the rewards that are available in that state and the amount of rewards that can be realized in states that are available from that state.

This may be done by either building a world model that compiles the reward function and the transition probabilities or omitting the use of a model. In the latter case, the agent obtains information about its environment by trial and error and computes estimates of the value of states or state-action pairs, in a way that estimates are cached.

Thus in model-free learning, by omitting the use of a model, cached state or state-action values are used and are updated upon subsequent learning. Conversely, predictions also concern the estimated values. The neurobiological substrate of model-free reinforcement learning is well rooted in the reward prediction error hypothesis of dopamine, i.e., upon encountering unexpected rewards or cues for unexpected rewards, ventral tegmental area (VTA) dopaminergic neurons burst fire. This phasic dopamine release is considered to be the manifestation of a reward prediction error signal computed as the difference between the expected and actual value of the reward received, and it drives model-free reinforcement learning.

Model-based learning, however is characterized by use of a world model, therefore direct experience is used to obtain the reward function and the transition

probabilities of the Bellman equation. Herein, learning is used to update the model (as opposed to model-free learning, where learning serves to update the cached estimated value of state). Generally, model-based reinforcement learning problems use the model to conduct forward-looking simulations for the sake of making predictions and/or optimizing policy in a way that the cumulated sum of the reward is maximized in the long term. This environmental model includes the set of state, the reward associated with each state, and the relationship between them and the probability of transitions between states that allow future simulations to be generated with states and state-action pairs.

These two distinct approaches to solve reinforcement learning problems are embodied by the concepts of model-based and model-free reinforcement learning, respectively. This distinction carries several implications about learning and updating the value of state as well as concerning the ability to carry out predictions, forward-looking simulations and optimization of behavior. Building on the theory of the ‘proactive brain’ and a related proactive framework that integrates model-free and model-based reinforcement learning, we expand the neurobiological foundations of model-based reinforcement learning. Previously, using the distinction for model-based and model-free learning and taking the structural and functional connectivity of neurobiological structures into consideration, we offered an overview of model-free and model-based structures.

1.3. The proactive model of the reinforcement learning

According to the proactive model of reinforcement learning (RL) the ventral striatum serves as a hub that anatomically connects model-free (pedunculo-pontine-tegmental nucleus (PPTgN) and VTA) and model-based [amygdala, hippocampus and orbitofrontal cortex (OFC)] structures, and integrates model-free and model-based inputs about rewards in a way that value is computed. Based on the structural connectivity of the ventral striatum and other, model-based structures (hippocampus, medial OFC (mOFC), amygdala), as well as their overlap with the default mode network, further suggested that the model used for model-based reinforcement learning

is built by the default mode network in a way that context frames are formed. By arranging the external world and internal states into context frames, information pertaining to valence and reward value of stimuli is reachable in such a way that it may assist making predictions facilitating action selection for maximizing reward.

Using the proactive brain concept we propose that OFC computes the reward function attribute of the Bellman equation, a function, that integrates state-reward contingencies and state-action-state' transactions (e.g. how executing an action determines transitioning from one state to the other one). Furthermore, using the proactive brain concept we suggest that the mOFC formulates reward expectations based on cue-context congruence by integrating cue (amygdala) and context (hippocampus) related input while the lateral OFC (lOFC) contributes to action selection by solving the credit assignment problem. Moreover we propose that anterior cingulate cortex (ACC) a key structure for action selection, computes the policy function of the Bellman equation by capturing reward history associated with previous action.

Additionally, using fundamental concepts of the proactive framework, we offer testable hypotheses based on the interaction between model-based and model-free systems. On one hand, we propose that the function of VTA dopaminergic neurons may be altered by manipulating OFC glutaminergic input. On the other, we propose that the model used by model-based reinforcement learning is updated by the interaction of the model-free and model-based accounts as model-free dopaminergic prediction error signals are able to influence the function of several model-based structures (OFC, hippocampus, amygdala, ACC, insular cortex).

1.4. Association between irisin/BDNF axis and brain reward system

Previous research by our group has described the relationship between the irisin/BDNF axis and the reinforcement system. Accordingly our group reported the alteration of the irisin/BDNF axis in mood disorders accompanying chronic obstructive pulmonary disease (COPD). Others have shown the association of BDNF gene polymorphism with anxiety and depression in asthma and thus suggested that genetic

variants may assume a significant role in the pathogenesis of distress disorder in the context of asthma.

Irisin is a contraction-regulated myokine, a highly conservative protein formed by proteolytic cleavage of fibronectin type III domain-containing protein 5 (FNDC5). Expression of FNDC5/irisin is most abundant in skeletal muscle and is enhanced by voluntary exercise. Upon its release into the systemic circulation, irisin may cause browning of white adipose tissue and is able to cross the blood-brain barrier. Moreover its expression is substantial in the central nervous system, e.g., midbrain and hippocampus, structures fundamentally involved in reinforcement learning. In the central nervous system (CNS) irisin can induce the expression of BDNF.

BDNF is a neurotrophin that is involved in preserving the plasticity of the brain by inducing long-term potentiation and consequent increase of the strength of synaptic connections inherent of reward-related changes and subsequent modified behavior. The link between mesocortico-limbic function and BDNF is well grounded as BDNF is expressed and has immediate influence on neuronal circuits activated by reinforcement learning and related reward-processing. BDNF as a growth factor enhances the viability of neurons, regulates the release of dopamine, and contributes to the so-called long-term potentiation, which is the basis for changes in reinforcement learning.

1.5. The relationship between the reinforcement learning and mood disturbance

The contextual learning is a basic mechanism involved in reinforcement (reward) learning and motivation. Several disorders have been mapped onto the reinforcement learning paradigm. Higher discounting rates (model-free parameter) for delayed rewards reflective of hopelessness and unwillingness was described in depression. It was further shown that anhedonia, is associated with diminished primary sensitivity to rewards (model-based parameter).

Based on findings of others, regarding altered contextual learning in distress disorder, a process inherent of reinforcement learning, as well as prior observations

pertaining to the potential involvement of BDNF in mood disorders and anxiety, assessing the presence and/or severity of distress disorder is feasible.

1.6. Mood disturbance, well-known co-morbid disorder of asthma

Generalized anxiety disorder and major depressive disorder, collectively termed distress disorder, are well-known co-morbidities of bronchial asthma. The significance of distress disorder in asthma symptom control, therapeutic adherence and quality of life is well-established and underscored. Increased prevalence of distress disorder in asthma patients in comparison with that in the general population is also recognized. Furthermore, a bidirectional relationship seems to exist between asthma and distress disorder, supported by large prospective studies showing robust association between mental disorders in childhood and subsequent evolution of adult onset asthma. Conversely, emergence of respiratory symptoms consistent with increased airway hyperreactivity or asthma exacerbations were reported in relation to even transient provocation of depressed or sad mood.

This two-way relationship suggests the possibility for a cause-and-effect relationship, and warrants the clarification of plausible shared neurobiological pathways underlying this link

1.7. Aims and goals

The prevailing change in behavior is driven by reward-related learning processes which may otherwise be responsible for the emergence and maintenance of harmful behaviors ("passions") in the background of major epidemics leading to premature death. It can be argued that these risk behaviors are at the root of dysfunction (or at least modification) of the mesocortical-limbic system, which can be attributed to the adverse change in the process of reinforcement learning as well. Based on these consideration, it is important to identify the neural structures in the background of behavioral changes, to map their functions, both from the neural and humoral aspects of regulation, which can potentially lead to the identification of new diagnostic and therapeutic targets.

Accordingly, we aimed at:

1. Identifying the structures and their functions involved in model-based reinforcement learning, particularly regarding the role orbitofrontal cortex assumes in the compilation of the Bellman equation's reward function.
2. Clarifying the clinical relevance of the proactive model of reinforcement learning.
3. Examining, in connection with the humoral regulation of the developed model, whether in case of asthmatic patients mood disturbance, due to the reinforcement learning disorder, is associated with changes in the irisin/BDNF axis.

2. Materials and methods

2.1. Conceptualization

On a theoretical basis, we developed a model that can explain how the brain creates the internal representations of the environment, how it identifies relevant stimuli, relationships, and actions. Building on the concept of the proactive model of reinforcement learning, the functional anatomical exploration of model-based learning, furthermore the proactive use of cue-context congruence, the role of the reinforcement learning's reward function was defined and potential clinical relevance was assessed. The model developed based on the theoretical considerations of this work is in line with the available experimental results.

2.2. Study Design and protocol

This is a cross-sectional study of asthma patients ($n = 163$) designed in compliance with the STROBE statement for cross-sectional studies. The present study adheres to the principles of the Declaration of Helsinki and was initiated after the approval of the Ethical Committee of the University of Debrecen (DEOEC RKEB/IKEB 3632-2012). Written informed consent was obtained from each participant prior to inclusion.

The recruitment period for the study was from September 1. 2012 to October 15. 2013. Consecutive patients attending the outpatient unit of the Department of Pulmonology (University of Debrecen) diagnosed with chronic inflammatory airway diseases (asthma, COPD, ACOS, and allergic rhinitis) were screened by pulmonologists blinded to the research hypothesis.

The current study assesses the potential involvement of altered irisin/BDNF axis in distress disorder among in our asthma cohort ($n = 167$). All patients participated in a control-based asthma management program complying with GINA as well as the relevant Hungarian practice guideline. Treatment at the time of inclusion was provided as clinically warranted hence most patients were regularly using inhaled corticosteroid preparations ($n = 156$) and/or long-acting β_2 agonists ($n = 141$). Exclusion criteria included history of malignant or benign tumors, or presence of acute inflammation of any kind during the preceding month. Demographic, anthropometric, and anamnestic (including history of smoking, diabetes, dyslipidemia, and hypertension) data were obtained. Smoking exposure was characterized by pack-years (that included past smoking and current smoking exposure).

Disease specific quality of life was characterized by using the official Hungarian version of the St. George's Respiratory Questionnaire (SGRQ) after obtaining written permission from its proprietor (Paul Jones, University of London, London, UK, on August 28. 2012).

2.3. Pulmonary function testing

Lung function was assessed using whole-body plethysmography, conducted in compliance with the ATS/ERS criteria (Piston whole-body plethysmograph PDT-111/p, Piston Medical, Hungary, equipped with automatic BTPS correction for cabin temperature, humidity and pressure as well as full automatic calibration and leakage test). Patients were asked to take their medication as usual on the morning of their examination (thus plethysmography was done with patients being on asthma control therapy). Best of three technically sound maneuvers was analyzed for each participant. At least two separate and technically appropriate measurements were conducted for resistance curves (each measurement consists recordings of at least 5 resistance loops) and results were accepted only if they were the same for both measurements. Raw pulmonary function data and percent predicted of normal reference values were recorded using algorithms provided by the manufacturer. The following parameters were included in the statistical analysis: Raw, FEV1% pred, FVC% pred, FEV1/FVC% pred, FEF25-75% % pred, RV% pred, TLC% pred, RV/TLC% pred.

2.4. Blood samples

Routine laboratory investigations were performed on the morning of the study after an overnight fast, according to the standard operational procedures of the Department of Laboratory Medicine (University of Debrecen). Briefly, serum or plasma samples were used to determine parameters descriptive of carbohydrate and lipid homeostasis, of kidney, liver, and muscle function, and of inflammation. CRP was dichotomized as high vs. normal using the cutoff of 4,6 and 5,2 mg/L for female and male patients, respectively. HOMA index was computed as previously. Serum samples for the determination of irisin and BDNF were frozen within 60 min and stored at -80°C until further analysis.

2.5. Determination of serum irisin and BDNF levels

Serum irisin levels were determined with a commercially available ELISA kit (Phoenix Pharmaceuticals, Burlingame, CA, USA), compliant of the manufacturer's instructions. Similarly, serum BDNF levels were measured in agreement with the manufacturer's instructions (Sigma-Aldrich, MO, USA). Absorbance was read at 450 nm an automated microplate reader. The irisin standard curve was linear from 1.34 to 29.0 ng/ml, and the detection limit was 1.34 ng/ml, in accordance with the information provided by the manufacturer. The detection limit for BDNF was <80 pg/ml. A standard curve showing linear relationship between optical density and concentration of irisin as well as BDNF were obtained with each plate. All measurements were performed in duplicate.

2.6. St. George's Respiratory Questionnaire (SGRQ)

The Hungarian version of SGRQ validated for a 1-month recall period was delivered using supervised self-administration, according to the SGRQ manual supplied by the copyright holder. SGRQ yields a Total score descriptive of the overall influence of the disease in terms of health status and three component scores, the Symptoms, the Activity, and the Impacts score. The main outcome measure of interest was the SGRQ Impacts score, as it characterizes disturbances of the psycho-social functioning, and thereby shows strong correlation with distress disorder. Scores are expressed as a percentage, with 100% representing the worst and 0% indicating the best possible subjective health status. Data entry was performed by two independent raters, who were blinded to patients' irisin and BDNF levels. Inter-rater variability assessed by Spearman correlation provided coefficients of 0.976 ($p < 0.001$), 0.997 ($p < 0.001$), 0.998 ($p < 0.001$), and 0.998 ($p < 0.001$) for the Symptoms, Activity, Impacts and Total scores, respectively. The mean of scores obtained by the independent raters was used for statistical analysis.

2.7. Statistical analysis

Normality for continuous variables was checked with Shapiro–Wilk test. If distribution was Gaussian, Student’s t-test was used for the comparison of two data sets (and data were presented as mean \pm SD), if not, Mann–Whitney U-test was performed (and data were presented as median; interquartile range). Frequencies were compared with Pearson’s χ^2 test.

Descriptive statistics were provided for the whole data set and the data dichotomized according to the median of two parameters, the Impacts score (as cutoff value). Patients with Impacts score $<22.53\%$ and $\geq 22.53\%$ were categorized into the lower ($n = 84$) and higher Impacts score group ($n = 83$), corresponding to the presence of less (or no) and more severe distress disorder, respectively.

The correlation between the Impacts score and (transformed) serum irisin level was established with Pearson’s correlation in the whole data set as well as in the lower and higher BDNF groups. Simple linear regression analysis was performed to identify possible determinants of the Impacts score (characteristic of distress disorder) as well as the (reciprocal of) serum irisin. The influence of traditional confounding factors (age, gender, height) as well as disease duration were assessed. Missing data were omitted. Simple linear regression revealed interaction between height and weight squared.

After the simple regression, age, gender, height (in interaction with weight squared) (as a priori variables) as well as all significant regressors were introduced into a multiple linear regression model to further quantify the relationship between the Impacts score and serum irisin concentration. Other SGRQ scores were omitted due to their high collinearity with the Impacts score. Lung function parameters yielding significant regressors were combined into a single parameter descriptive of lung function using principal component analysis. Variables were entered in the model simultaneously, and then factors not significantly contributing to the model were deleted. (Eventually, the final model contained all variables identified a priori, the composite of lung function parameters and the history of dyslipidemia.)

In addition, the final model was stratified with respect to the median of BDNF. Patients with BDNF level $<311,4 \mu\text{mol/L}$ and $\geq 311,4 \mu\text{mol/L}$ were categorized into the

lower (n = 84) and higher BDNF group (n = 83), respectively. Heteroskedasticity and goodness of fit for the model was assessed by Cook-Weisberg and Ramsey tests, respectively.

Statistical analysis was performed with Stata 13.0 software (Stata Corporation). Values are given as mean \pm SD or medians (with interquartile ranges: IQR), excepting regression coefficients which are presented with their 95% confidence intervals (CI).

3. Results

3.1. The proactive brain builds a model of the environment

A key issue of model-based learning concerns to how the brain creates the internal representations of the environment, thus how it segments and identifies relevant stimuli, contexts and actions. The world model must represent the salient features of the external and internal (interoceptive, viscerosensory, affective and cognitive) environment. Previously, building on the proactive brain concept, we have proposed that model-based learning utilizes association-based context frames to build its world model, upon which forward looking mental simulations and predictions may be formulated. A key to this concept is the creation of context frames. This is done by arranging stimuli (e.g. unconditioned stimuli and their conditioned cues) and their contexts into context frames. Contexts encompass internal [cognitive/affective (including reward-related), interoceptive (physiological and neurohumoral)] and external (spatial, temporal, social or cultural) settings, thus context frames contain a priori information about the scalar value of reward.

Furthermore, context frames come to signal cue-context associations reflecting statistical regularities and a lifetime of extracting patterns from the environment (related to contingencies, spatial locations, temporal integration, etc.). Organization of context frames enables rudimentary cue- or context-related information to retrieve the most relevant context frame from memory, by means of associative processes. Furthermore it helps to cope with ambiguity and uncertainty, as coarse contextual information is sufficient to activate the most relevant context frame, which may assist in predicting the most probable identity of the cue.

The environment is transformed into context frames by means of cue and context conditioning. Cue and context conditioning are two concepts familiar to Pavlovian learning, with cue conditioning being the central paradigm. Cue and context conditioning are done by parallel but richly interconnected systems, with prior research pinpointing the amygdala as a neural substrate that is the prerequisite for affective processing of a stimuli as well as for cue-conditioning (e.g. forming associations

between cues and primary reinforcers). Furthermore, amygdaloid input, representing subcortical inferences pertaining to the affective and motivational value of the stimulus, is incorporated into decisions by function of the OFC. Hippocampus assumes a central role in context conditioning, as the hippocampal area is critical for providing complex representation of signals; and its link with the OFC has been implicated in the integration of declarative representations with other information to guide behavior. Additionally, recent observations showed an interaction between the hippocampus and OFC in support of context-guided memory. Furthermore using this proactive framework, we have previously proposed that the basolateral amygdala computes cue-reward, while the hippocampus forms context-reward contingencies, respectively.

3.2. The orbitofrontal cortex compounds the reward function attribute of the Bellmann equation

The Bellman equation descriptive of the agent's knowledge of the environment, is built by the OFC with its distinct parts assuming well differentiated roles (the medial and lateral part contributing to state-reward contingency and state-action-state contingencies, respectively). The medial OFC (mOFC) integrates cue- and context-based pieces of information provided by the amygdala and hippocampus, respectively, to assess cue-context congruence. Based on cue-context congruence, it identifies the context frame most relevant for a given state, to extract information regarding reward expectations. The lOFC may contribute to the credit assignment domain of action selection by having access to information about state- action-state' contingencies. The integrative function of OFC is well in agreement with its anatomical position, as it complies input from all sensory (e.g. visual, auditory etc.) modalities and subcortical (e.g. hippocampus, amygdala, ventral striatum, VTA, etc.) areas. In line with this central position is OFC's ability to integrate concrete and abstract multisensory perceptual input with memories about previous stimuli, state transactions as well as affective and incentive value of associated outcomes. OFC ensembles encompass information about context, stimuli, behavioral responses as well as rewards associated with states. The significance of OFC in integrating contextual information into decisions.

OFC's contribution to the other key element of the reward function, e.g. credit assignment. Credit assignment, one of the two domains determining action selection, is the association of behaviorally relevant stimulus with the action leading to preferable outcomes, by detecting state-action-state' contingencies.

The OFC conjointly signals information about reward identity, value, location, behavioral responses and other features was corroborated by works showing that OFC neurons encode all aspects of a task, they attribute rewards to preceding states and code state transitions. This underscores OFC's role in adapting to changing environments by enabling flexible behavior facilitated by the formation of new associations between cues (states), state transitions and rewards via indirect links with other brain areas. The critical contribution of OFC in influencing ongoing behavior and updating associative information.

In summary, the role of OFC is crucial in reinforcement learning, as it determines the attribute of the reinforced learning reward, thus contributing to model-based reinforcement learning through its integrative function. From the amygdala (where subcortical processing of the stimulus' emotional and motivational value takes place) and the integration of information from the hippocampus (which generates contextual reward contingents), mOFC allows contextual frameworks to be established by means of signal-context congruence evaluation and access to state-action-state contingents provision. These contextual frameworks also include information related to the expected reward, which influences motivated behavior according to the proactive model of reinforcement learning, enabling the interpretation of the mechanism of action of cognitive psychotherapy.

3.3. Association between serum levels of irisin and BDNF and mood disturbances in asthma patients

3.3.1. Patients

The basic characteristics of our asthma patient cohort (n = 167) were as follows. The mean age of the patient population is 48 years, 91 patients were female and 22 patients were smokers. The median irisin level (IQR) was 7.87 (7.15–8.82) ng/mL and the mean of serum \pm BDNF level was 314.46 ± 118.68 ng/mL in the whole sample. No statistical difference was present in either serum irisin or serum BDNF levels with respect to gender.

3.3.2. Comparison of patients regarding distress disorder

The two strata of our asthma cohort, dichotomized based on the median of Impacts score, were homogenous regarding most parameters investigated. Nonetheless, patients with more severe form of distress disorder were significantly older, shorter, had smoked more, and had higher serum triglyceride and fibrinogen but lower irisin levels. In addition, they performed poorer in the lung function test reflected by worse parameters in terms of FVC% pred, FEV1% pred, FEF25-75% pred %, RV% pred, RV/TLC% pred, and Raw, and had significantly higher Symptoms, Activity and (as a consequence) Total scores. Furthermore, higher proportion of patients in the higher Impacts score group suffered from hypertension and dyslipidemia.

3.3.3. Association between SGRQ's Impacts score, serum irisin and BDNF levels

The Impacts score and reciprocal of serum irisin level showed significant positive correlation in the whole patient cohort (Pearson correlation coefficient: 0.19, $p = 0.014$). The correlation coefficient increased indicating stronger correlation (and, despite the smaller patient number, remained statistically significant) in the stratum with higher BDNF levels, while it decreased (and lost its statistical significance) in the lower BDNF group (Pearson correlation coefficient: 0.25 and 0.11, $p = 0.025$ and $p = 0.30$,

respectively). This suggests that serum irisin level is associated with processes that lead to disease-related impairment of psycho-social function among asthma patients, an effect that is more pronounced if higher serum irisin concentration is accompanied with higher serum BDNF level.

Simple linear regression yielded corroborating results upon evaluating the relationship between Impacts score and reciprocal of irisin. This relationship remained significant after adjusting for all significant predictors and a priori determinants by means of multiple linear regression. The Impacts score and reciprocal of irisin showed a strong, significant, positive association (β : 147.74; CI: 42.17, 253.30; $p = 0.006$), which was even more pronounced (e.g., the regression coefficient was higher) in the higher BDNF group (β : 213.38; CI: 56.38, 370.38; $p = 0.008$), while it was weaker (an effect reflected by lower regression coefficient and lack of statistical significance) in the lower BDNF group (β : 60.89; CI: -108.55, 230.33; $p = 0.48$). All models (the full model as well as the stratified ones) were significant ($p < 0.001$, $p = 0.040$, $p = 0.003$). The Cook-Weisberg test showed no heteroskedasticity for the full model and strata with higher and lower BDNF levels ($p = 0.55$, $p = 0.22$, and $p = 0.79$, respectively). Moreover, all three models showed good fit reflected by the locally weighted scatterplot smoothing as well as by the Ramsey test ($p = 0.51$, $p = 0.28$, and $p = 0.55$ for the whole data set and strata with higher and lower BDNF concentrations, respectively).

The final multiple linear regression model (built for the Impacts score) contained only dyslipidemia in addition to the reciprocal of irisin and a priori identified parameters. The presence of dyslipidemia showed a significant, positive association with the Impacts score of asthma patients (β : 7.01; CI: 0.85, 13.35; $p = 0.026$), a relationship that was only significant in the whole data set.

4. Discussion

4.1. 'Proactive' use of cue-context congruence

4.1.1. Specificity of the OFC regarding the reward function

The OFC represents models for reinforcement learning, our proposition furthers this concept by linking a specific attribute of the Bellman equation descriptive of reinforcement learning to OFC function. A key new finding concerns the use of cue-context associations (deducted from the proactive brain concept) to explain OFC's integrative function, with respect to cue- and context-related inputs (coming from the amygdala and hippocampus, respectively), reward expectations and credit assignment. Therefore we propose that the OFC computes the reward function attribute of the Bellman equation and thereby contributes to model-based reinforcement learning by assessing cue-context congruence along and maps cue/context/action-reward contingencies to context frames. By using the reward function, the OFC is able to signal predictions related to reward expectation. OFC's role in reinforcement learning guided decision making concerns the ability to make detailed, exible and adjustable predictions on context frames modelling the environment by assessing cue-context congruence and by means of credit assignment.

To assess the specificity of our model we overviewed the function of other, significant interconnected structures implied in contributing to reinforcement learning, e.g. ACC, dorsolateral prefrontal cortex (dlPFC), pre-supplementary motor cortex (preSMC) and insular cortex.

With respect to ACC, its most commonly agreed upon feature is its engagement in decision making tasks that demand cognitive control. ACC is involved in action selection based on the assessment of action-outcome relations. Conversely it is involved in monitoring and integrating the outcome of actions. In addition to governing the relationship between previous action history and next action choice, the ACC assumes a complementary role in exploratory generation of new action for the action set, used by reinforcement learning. Summarizing, ACC is involved in one of the two domains of action selection.

Insular cortex's contribution may be assessed in terms of model-based reinforcement learning, given its dense connections with model-based structures including amygdala nuclei, OFC, ventral striatum, ACC and the dlPFC. In line with its functional connectivity, insula is responsible for detecting behaviorally salient stimuli and coordination of neural resources. By means of its anatomical connections insula is able to integrate ascending interoceptive and viscerosensory inputs in a way that subjective feelings are transformed to salience signals influential of decision making. The insula has a central role in salience processing across several domains and is involved in mediating the switching between the activation of the default mode network and the executive network to ensure optimal response to salient stimuli thus confers indirect, yet significant influence on model-based reinforcement learning.

4.1.2. Computation of the model-free reward prediction error using input from the OFC

Several testable hypotheses come from the bidirectional interactions between model-free and model-based learning. The OFC is known to project glutaminergic efferents to several structures involved in model-free reward prediction error signaling, including the PPTgN (that offers one of the strongest excitatory drives to the VTA), VTA (that emits the model-free dopamine learning signal) and ventral striatum (that is responsible for computing value by compounding varying inputs). By reaching PPTgN, OFC may modulate the VTA's most significant stimulating afferent, while OFC's influence on dopaminergic neurons of VTA can extend to the alteration of both the spike and burst activity of dopaminergic neurons. The OFC's reward expectation signal contributes to the detection of error in the reward prediction error signal, if contingencies are changing.

Furthermore, expectancy-related changes in firing of dopamine neurons were shown to hinge on orbitofrontal input. This led to the conclusion that the OFC's contribution to prediction errors is via its influence on dopamine neurons. The function of VTA dopaminergic neurons may be altered by cue-context manipulations leading to

the change of glutaminergic input emanating from OFC, or by other interventions like transcranial magnetic stimulation.

4.1.3. Updating the world model

Linking our proactive model of reinforcement learning to the mathematical formalism of the Bellman equation gives a framework to jointly draw inferences concerning spatiotemporal environmental contingencies included in the reward function and action selection reflective of the reward structure contributing policy formation. Information about the scalar value of reward is encoded in context frames based on its spatiotemporal proximity with cues. This is done in a way that context frames may be mobilized based on cue-context congruence. Nonetheless it may be further inferred from our proactive model that feedback regarding the scalar value of reward, signaled as reward prediction error, may update the reward attribute of the cue-relevant context frame as follows.

The main targets of VTA dopaminergic neurons are the ventral striatum, amygdala, hippocampus, OFC, ACC and insular cortex. Considering the three factor rule, an extended form of the Hebbian rule, i.e. synaptic strength is increased if the simultaneous presynaptic and postsynaptic excitation coincides with dopamine release by means of long-term potentiation, it may be postulated that in the event of dopamine release cue (amygdala), context (hippocampus) and cue-context congruence (OFC) relations are wired together, thus altering the reward structure (e.g. the environmental model). Therefore, the model-free reward prediction error output is necessary for updating the world model subserving the model-based system.

Summarizing, this implication offers further indirect support for the interaction between model-free and model-based accounts by suggesting that model-free reward prediction error signal may contribute to updating the model used by model-based learning by altering the scalar value of rewards in the relevant context frames and it updates the policy underlying action selection to maximize outcomes.

4.1.4. Clinical implications of our hypothesis

The theoretical collision of the concept of proactive brain with that of reinforcement learning has substantial clinical relevance. The reward-related information and action selection is governed by cue and context information, we offer a framework for behavior modification. Given that reward information used by reinforcement learning depends on the statistical regularities of cue-context-reward co-occurrence, direct manipulation of cue-context-reward contingencies could overwrite former regularities to alter the reward function. Some currently used techniques of cognitive behavioral therapy (e.g. desensitization, chaining, triple or seven column technique) could be interpreted in terms of this framework. Furthermore, exploitation of technological advancements could be used to facilitate mental processes such as daydreaming or visualization that contribute to the alteration of the model used by model-based learning. With the help of current technology, patients engage in activities in virtual settings, facing experiences that, according to our concept, would serve as input for shaping future behavior by formation of novel Pavlovian learning-based associations that alter existing spatio-temporal contiguities of cues, contexts and rewards, and may even extend to changes in state-state' transitions.

4.1.5. The relationship between irisin/BDNF axis and mood in asthma patients

Our major finding is that severity of distress disorder showed robust positive linear relationship with the reciprocal of serum irisin among adult asthma patients, undertaking asthma controller therapy. Interestingly, severity of distress disorder was more closely associated with the (reciprocal of) serum irisin concentration in the patient stratum with higher BDNF levels, while this association became statistically nonsignificant in the stratum with lower BDNF levels. In addition, the final model identified the presence of dyslipidemia as a significant regressor influencing the Impacts score.

Our main finding, that the positive linear relationship between the Impacts score reflective of distress disorder and the (reciprocal of) serum irisin varies with respect to

the BDNF level, implicates the possibility of an interaction between these two factors in the context of distress disorder. This finding corroborates our previous hypothesis that serum irisin, via BDNF, may influence reinforcement learning and allied processes including contextual learning.

Convergence of previously distinct entities (anxiety disorder and major depressive disorder) in such phenomenological nosological systems is well aligned with the need of identifying common neurobiological pathways involved in the etiopathogenesis of distress disorder. Linking psychiatric disorders (depressive disorders) to the theoretical framework of the proactive model of reinforcement learning acknowledges the possibility that these disorders may evolve due to the disruption of the same core system.

Reinforcement learning is a central process that guides individual behavior in a way that immediate and cumulated future rewards are maximized in the long term. The brain has two fundamentally different, but closely interconnected systems governing model-based and model-free reinforcement learning, paradigms that differ in terms of use of a world model. The mutual influence of these systems has been implicated previously as the dopaminergic reward prediction error signal may contribute to updating the reward attribute of the context frames (by means of VTA's dopaminergic input on the hippocampus, amygdala and mOFC), furthermore glutaminergic outflow of mOFC modulates the function of model-free structures including VTA.

Irisin and its downstream mediator BDNF are humoral factors that target several structures and, accordingly, may modulate reinforcement learning. Irisin, either of peripheral or neuronal origin, has the ability to trigger the expression of BDNF in these structures. BDNF regulates dopamine 3 (D3) signaling pathways in VTA, a pathway that has fundamental influence on dopamine's effect with respect to structural plasticity in the model-free reinforcement learning system. Furthermore, high expression of BDNF parallels enhancement of long-term potentiation underlying synaptic plasticity, neurogenesis, and neuronal differentiation in the model-based system's amygdala, hippocampus, and prefrontal cortex. Summarizing, the irisin/BDNF axis modulates several structures constituting the neurobiological system underlying reinforcement learning.

In our previous work, we investigated the possible involvement of the irisin/BDNF axis in the evolution of distress disorder in a COPD patient cohort. Although, similarly to our current findings in asthma, a significant linear relationship was present between the Impacts score and reciprocal of serum irisin levels among COPD patients, the interaction between serum irisin and BDNF levels was found to be different. In the case of COPD patients, the linear association became stronger in the stratum of patients with BDNF levels in the lower half. On contrary, results in our cohort of asthma patients in the present study showed more robust linear association in the stratum comprising patients in the upper half of the BDNF levels. Several causes may underlie this difference. On one hand, the different role of BDNF in the etiopathogenesis of these two similar, but distinct disease entities must be acknowledged. The ability of BDNF to induce airflow limitation and to promote airway hyperresponsiveness to histamine has been described in animal models of asthma as well as in asthma patients. Meanwhile in COPD, BDNF's role in airway pathology was only implicated in the context of airway remodeling. Prior clinical reports have shown the ability of β 2 agonists to increase BDNF concentration in platelets and serum, parallel to the deterioration of airway responsiveness. This deleterious effect may be counterbalanced by co-administration of inhaled corticosteroids. Based on these considerations, differential alteration of BDNF-related mechanisms in these two diseases seems plausible, although further investigations are needed.

In the current study, based on the final multiple regression model, we have found that dyslipidemia is a significant determinant of the Impacts score, indicating that disruption of lipid metabolism may worsen the severity of distress disorder in asthma. Dyslipidemia is known to be associated with increased asthma prevalence, moreover, it has been implicated as a risk factor for asthma.

Taking these considerations into account, it may be suggested that changes in the irisin/BDNF axis may lead to impairment of reward-related motivation, hindering the patient's adherence to treating asthma and thus reducing the effectiveness of the treatment.

5. New findings

1. The integrative function of orbitofrontal cortex (OFC) is crucial in the reinforcement learning decision-making:
 - the medial OFC identifies contextual frames based on cue-context congruence
 - the lateral OFC provides access to the action selection interval
2. Anterior Cingular Cortex (ACC) has an important role in action selection by recording the reward history associated with the previous actions.
3. The proactive model of reinforcement learning provides an opportunity to interpret the mechanism of action of cognitive psychotherapy.
4. Distress disorders, known to be associated with altered contextual learning (a funamental process underlying reinforcement learning) are linked with changes of thethe irisin/BDNF axis in asthmatic patients.

6. Summary

In summary, we put forward several testable hypotheses. We postulated several structures of the model-based network to be involved in computing specific attributes of the Bellman equation, the mathematical formalism used to conceptualize machine learning based accounts of reinforcement learning. Furthermore, we provided a plausible mechanism of how the model, used by model-based learning system, is created by organizing cue, context, reward information into context frames and capturing conjoint information of stimulus, action and reward. In addition, based on the bidirectional interaction of model-free and model-based structures we made two further propositions. One, given the reward value related input to the model-free structures (PPTgN and VTA), cue-context manipulations or transcranial magnetic stimulation may be applied to alter the model-free dopaminergic signal. Two, reward prediction error related dopamine signal may contribute to the update of both the model and the policy functions of model-based reinforcement learning. Finally, our proactive framework for reinforcement learning has clinical implications as it builds on the use of cue-context associations to offer a representational architecture, upon which behavioral interventions may be conceptualized.

In addition to our theoretical work, we attested the hypothesis that altered reinforcement learning based processes e.g. contextual learning manifested clinically in the form of distress disorders is paralleled by the alteration of the irisin/BDNF axis in asthma patients. To the best of our knowledge, the current study is the first to assess the relationship among serum irisin level, serum BDNF level and the occurrence (and severity) of distress disorder in asthma. The strong positive association between the reciprocal of serum irisin and Impacts score (i.e., the inverse proportionality between serum irisin and Impacts score) in the multiple linear regression model indicates that circulating irisin is associated with a lower occurrence and/or milder extent of distress disorder. Our further finding that irisin's effect is more pronounced if serum BDNF level is higher adds clinical evidence to prior observations in preclinical models regarding this pathway. Summarizing, irisin may contribute to BDNF's ability to influence mood and anxiety by altering contextual learning in structures under BDNF control thus induce the evolution of distress disorder.

7. Acknowledgements

This research was supported by the Hungarian Brain Research Program (**KTIA_13_NAP-A-V/2**), the Scientific Research Grant of the Hungarian Foundation for Pulmonology (awarded in 2015), the Ministry for National Economy (**GINOP-2.3.2-15- 2016-00062 and AGR-PIAC-13-1-2013-0008**). Writing the thesis was supported by the Ministry for National Economy (**GINOP-2.3.2-15-2016-00005**).



Registry number: DEENK/43/2018.PL
Subject: PhD Publikációs Lista

Candidate: Magdolna Emma Szilasi
Neptun ID: QR9KUG
Doctoral School: Doctoral School of Health Sciences

List of publications related to the dissertation

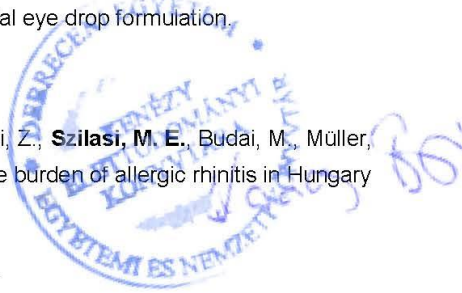
1. **Szilasi, M. E.**, Pák, K., Kardos, L., Varga, V. E., Seres, I., Mikáczó, A., Fodor, A., Szilasi, M., Tajti, G., Papp, C., Gesztelyi, R., Zsuga, J.: The alteration of irisin - brain-derived neurotrophic factor axis parallels severity of distress disorder in bronchial asthma patients.
Front. Neurosci. 11, 1-12, 2017.
DOI: <http://dx.doi.org/10.3389/fnins.2017.00653>
IF: 3.566 (2016)
2. Zsuga, J., Bíró, K., Tajti, G., **Szilasi, M. E.**, Papp, C., Juhász, B., Gesztelyi, R.: 'Proactive' use of cue-context congruence for building reinforcement learning's reward function.
BMC Neurosci. 17 (70), 70, 2016.
DOI: <http://dx.doi.org/10.1186/s12868-016-0302-7>
IF: 2.312





List of other publications

3. Tajti, G., Papp, C., Kardos, L., Kéki, S., Pák, K., **Szilasi, M. E.**, Gesztelyi, R., Mikáczó, A., Fodor, A., Szilasi, M., Zsuga, J.: Positive correlation of airway resistance and serum asymmetric dimethylarginine (ADMA) in bronchial asthma patients lacking evidence for systemic inflammation.
Allerg Asthma Clin. Immunol. 14 (2), 1-12 -, 2018.
IF: 2.869 (2016)
4. Tajti, G., Gesztelyi, R., Pák, K., Papp, C., Kéki, S., **Szilasi, M. E.**, Mikáczó, A., Fodor, A., Szilasi, M., Zsuga, J.: Positive correlation of airway resistance and serum asymmetric dimethylarginine level in COPD patients with systemic markers of low-grade inflammation.
Int. J. Chronic Obstr. Pulm. Dis. 12, 873-884, 2017.
DOI: <http://dx.doi.org/10.2147/COPD.S127373>
IF: 3.157 (2016)
5. Ivancsó, I., Bohács, A., Szalay, B., Toldi, G., **Szilasi, M. E.**, Müller, V., Losonczy, G., Rigó, J., Vásárhelyi, B., Tamási, L.: Circulating periostin level in asthmatic pregnancy.
J. Asthma. 53 (9), 900-906, 2016.
DOI: <http://dx.doi.org/10.3109/02770903.2016.1165697>
IF: 1.746
6. Roy, R. J., Budai, L., Kiss, L., **Szilasi, M. E.**, Petrikovics, I., Budai, M.: Liposomal Encapsulation of Photoprotector Molecules.
J. Pharm. Nanotech. 3 (3), 18-21, 2015.
7. Csillag, A., Kumar, B. V., Szabó, K., Szilasi, M., Papp, Z., **Szilasi, M. E.**, Pázmándi, K. L., Boldogh, I., Rajnavölgyi, É., Bácsi, A., László, F. J.: Exposure to inhomogeneous static magnetic field beneficially affects allergic inflammation in a murine model.
J. R. Soc. Interface. 11 (95), 20140097-20140097, 2014.
DOI: <http://dx.doi.org/10.1098/rsif.2014.0097>
IF: 3.917
8. Wijesooriya, C., Budai, M., Budai, L., **Szilasi, M. E.**, Petrikovics, I.: Optimization of liposomal encapsulation for ceftazidime for developing a potential eye drop formulation.
J. Basic Clin. Pharma. 4 (3), 73-75, 2013.
DOI: <http://dx.doi.org/10.4103/0976-0105.118810>
9. Szilasi, M., Gálffy, G., Fónay, K., Márk, Z., Rónai, Z., Szalai, Z., **Szilasi, M. E.**, Budai, M., Müller, V., Somfay, A., Horváth, I., Tamási, L.: A survey of the burden of allergic rhinitis in Hungary from a specialist's perspective.
Multidiscip. Respir. Med. 7 (1), 1-6, 2012.
DOI: [http://dx.doi.org/10.1016/0047-6374\(85\)90041-7](http://dx.doi.org/10.1016/0047-6374(85)90041-7)
IF: 0.151





10. Szilasi, M., Fodor, A., Szluha, K., Szántó, J., Tóth, L., **Szilasi, M. E.**, Budai, M., Adamecz, Z.:
Ismeretlen primer tumor kemo-radioterápiás kezelésnek eredményei.
Med. Thorac. 65 (2), 106-110, 2012.
11. Szilasi, M., Horváth, Á., Dolinay, T., Szántó, J., Brugós, L., Lengyel, L., Kiss, S. S., **Szilasi, M. E.**,
Adamecz, Z.: IIIA-B stádiumú NSCLC elsővonalbeli Taxotere-Cisplatin konkuráló kemo-
radioterápiájának eredményei 42 irreszekábilis daganatos beteg vizsgálata alapján.
Magyar Onkol. 50 (3), 233-236, 2006.

Total IF of journals (all publications): 17,718

Total IF of journals (publications related to the dissertation): 5,878

The Candidate's publication data submitted to the iDEa Tudóstér have been validated by DEENK on the basis of Web of Science, Scopus and Journal Citation Report (Impact Factor) databases.

20 February, 2018

